Non-parametric resampling of random walks for spectral network clustering

Fabrizio De Vico Fallani,[1] Vincenzo Nicosia,[2] Vito Latora,[2,3] and Mario Chavez[1]

[1] *CNRS UMR-7225, Hôpital de la Pitié-Salpêtrière. Paris, France*
[2] *School of Mathematical Sciences, Queen Mary University of London, Mile End Road, E1 4NS, London (UK).*
[3] *Dipartimento di Fisica e Astronomia, Universitá di Catania, Via S. Sofia 61, 95123, Catania (Italy)*

(Dated: April 16, 2013)

Parametric resampling schemes have been recently employed in complex network analysis to assess the statistical significance of graph clustering and the robustness of community partitions. Here we propose a method to detect significant communities in complex networks based on the non-parametric resampling of the transition matrix associated with an unbiased random walk on the graph. We test this non-parametric approach on real-world and synthetic modular networks and we show that it substantially improves the existing algorithms based on spectral network decomposition. We finally discuss how bootstrap distributions can be used to assess the significance of different network properties without any a-priori assumption about the structure of the ensemble to which the network belongs.

PACS numbers: 89.75.-k, 02.50.Ga, 05.10.Ln

In the past decade, network science has proven to be a robust and comprehensive framework to investigate, model and understand the structure and function of the complex interaction patterns observed in diverse biological, physical, social and technological systems [1–3]. One of the central problems in the characterization of complex networks is the identification of communities, i.e. tightly-knit groups of nodes which exhibit poor connectivity with the rest of the graph [4]. In fact, experimental evidences confirmed that communities are the meso-scale building blocks of complex networks: they usually correspond to functional modules in the brain [5, 6], to topical clusters in social and communication networks [7], to metabolic reactions and functional domains in protein interaction networks [8, 9], to disciplines and research areas in collaboration networks [4]. For this reason, many different community detection algorithms have been proposed [10].

A typical problem of networks analysis is that each network represents a single observation drawn from an unknown distribution of graphs having a certain set of characteristics [11]. Consequently, there is no predefined way to assess the statistical significance of any metric estimated on a given network.

A widely approach to estimate the statistical significance of an observed network property takes into account *randomized network ensembles*, i.e. sets of graphs obtained from the original network by keeping fixed some structural properties (e.g. the degree sequence or the clustering coefficient) and rewiring the edges at random [12–14]. In the case of community detection, this approach led to the definition of the modularity function, which quantifies the significance of a given community partition of a graph as the deviation from the average modularity expected in an ensemble of random graphs having the same degree sequence [7].

Another possibility is *parametric bootstrapping*, in which the significance of a network property is assessed against small perturbations of the graph connectivity [15–17]. This approach relies on the hypothesis that the observed network is representative of a set of graphs (a model) having a certain (*a priori* known) distribution of structural characteristics. Consequently, the significance of an observed network feature is estimated as the deviation from the average of the corresponding model. Many different parametric resampling schemes have been used to assess the robustness of the modular networks against small perturbations of the connectivity. However, all these methods require an *a priori* hypothesis about the model to which the network belongs, so that the unbiased statistical assessment of a network partition remains an open challenge [18].

A solution which does not require any *a priory* knowledge is *non-parametric bootstrapping*, a computer-based technique for providing the statistical confidence of almost any statistical estimate [19, 20]. The key principle of non-parametric bootstrapping is to simulate repeated observations from an unknown population using the available data samples (in our case, a single network) as a starting point.

In this Brief Report we propose a method to identify statistically significant community partitions using non-parametric bootstrapping. The method is based on the construction of replicates of the transition matrix of the network, and in estimating an average distance matrix whose elements represent the expected spectral distances between pairs of nodes of the graph. Then, the obtained distance matrix is fed into a hierarchical clustering algorithm. By using non-parametric bootstrapping we finally construct a community partition for a graph by combining the community partitions obtained on the different bootstrap replicas of the same graph. This approach is in the same line of ensemble or consensus clustering methods, which combine several partitions generated by different clustering algorithms —or by different runs of the same algorithm— into a single statistically significant partition [16, 17, 21–24]. We analyze the community partitions obtained by non-parametric bootstrapping on different synthetic and real-world modular networks, and

we show that this approach can substantially improve the performances of existing spectral clustering methods. Moreover, we illustrate how non-parametric bootstrap can be also used to assess the significance of other structural properties, e.g. the spectral distance among nodes.

Bootstrapping random walks on graphs.– Let us consider a connected undirected and unweighted graph $G(V, E)$ with $N = |V|$ nodes and $K = |E|$ edges, associated to the binary adjacency matrix $A = \{a_{ij}\}$ ($a_{ij} = 1$ if there is an edge connecting node i and node j, while $a_{ij} = 0$ otherwise). We define an unbiased random walk on the graph G as the time invariant ergodic Markov chain whose transition matrix reads $\mathbf{P} = \{P_{ij}\}$, where $P_{ij} = a_{ij}/k_i$ is the probability for a walker on node i to jump, in one time step, from node i to one of its neighbours j (here we denote by $k_i = \sum_j a_{ij}$ the degree of node i). If we call $Q(t) = \{q_i(t)\}$ the vector of the probabilities to find a walker on each node i at time t, such that $\sum_i q_i(t) = 1\,\forall t$, the time evolution of the random walk is given by the equation:

$$Q(t+1) = \mathbf{P}^\intercal Q(t) \tag{1}$$

If the adjacency matrix A is primitive, the Perron–Frobenius theorem guarantees that Eq. (1) converges in finite time to a unique stationary state distribution $Q^* = \{q_i^*\}$, such that $Q^* = \mathbf{P}^\intercal Q^*$ [25, 26].

The authors of Ref. [27] proposed a generic bootstrap scheme to resample the transition probabilities of a finite state time-invariant Markov chain. Starting from a realization χ of the Markov chain, one constructs the maximum likelihood estimator of the associated transition matrix $\mathbf{P}$ as $P_{ij} = \frac{f_{ij}}{f_i}$, where f_{ij} the observed number of transitions from state i to state j in χ and $f_i = \sum_j f_{ij}$. Then, replicates of the observed transition matrix are obtained by drawing the random variables $\{f_{i1}^*, \ldots, f_{iN}^*\} \sim \text{Multinomial}(f_i; p_{i1}, \ldots, p_{iN})$ according to $\widetilde{P_{ij}} = \frac{f_{ij}^*}{f_i}$. The distribution of $\widetilde{\mathbf{P}}$ is then obtained by Monte-Carlo sampling. This approach was shown to be asymptotically valid for approximating the sampling distribution of $\mathbf{P}$ [27], and has been also used to assess the confidence intervals of transition probabilities in disease modeling [28].

Since the unbiased random walk on the graph G defined by Eq. (1) is a finite-state time-invariant Markov chain, we can construct a similar resampling scheme in which replicates of the transition matrix $\mathbf{P}$ are obtained by randomly drawing the variables $\{f_{i1}^*, \ldots, f_{iN}^*\}$ from a multinomial distribution with probabilities $\{p_{i1}, \ldots, p_{iN}\}$, conditional on the observed number of edges k_i. Each replicate of the original transition matrix is estimated as $\widetilde{P_{ij}} = \frac{f_{ij}^*}{k_i}$. It is worth noticing that, in contrast to previous approaches where each link was resampled independently from the others, here the replicas of the transition probabilities for each node i are drawn from a multinomial distribution, accounting for the observed transitions to other nodes $\{p_{i1}, \ldots, p_{iN}\}$.

Spectral clustering with bootstrap information.– Spectral clustering is a widely used technique to identify clusters and communities in a graph. It consists in projecting each node i of G into a point x_i of an appropriate metric space $\mathbf{X}$. The coordinates of x_i in $\mathbf{X}$ are given by the components associated to node i of the eigenvectors of the adjacency matrix A of G or, more frequently, of the transition matrix $\mathbf{P}$ [5, 29]. The mapping of G on the space $\mathbf{X}$ allows to cluster the nodes according to their Euclidean distance in $\mathbf{X}$, so that nodes whose projections in $\mathbf{X}$ are closer have a higher probability to be put in the same cluster.

The transition matrix $\mathbf{P}$ of the random walk on G is characterized by a set of eigenvalues $\{\lambda_0, \lambda_1, \ldots, \lambda_{N-1}\}$ such that $|\lambda_0| \geq |\lambda_1| \geq \ldots \geq |\lambda_{N-1}|$. Each eigenvalue λ_k is associated to the left and right eigenvectors φ_k and ψ_k, which satisfy $\varphi_k^\intercal \mathbf{P} = \lambda_k \varphi_k^\intercal$ and $\mathbf{P}\psi_k = \lambda_k \psi_k$, respectively. The mapping of the matrix $\mathbf{P}$ on an Euclidean space produces a geometric diffusion of points such that the distance between the points x_i and x_j, corresponding to the nodes i and j, can be written as [30]: $d_{ij}^2 = \sum_{k\geq 1} \lambda_k^2 (\psi_k(i) - \psi_k(j))^2$ where $\psi_k(j)$ denotes the j^{th} component of the k^{th} right eigenvector (notice that $\varphi_0 = Q*$ and ψ_0 is a constant vector). This distance can be approximated by using only the first β nontrivial eigenvectors and eigenvalues:

$$d_{ij}^2 \simeq \sum_{n=1}^{\beta} \lambda_n^2 (\psi_n(i) - \psi_n(j))^2. \tag{2}$$

This mapping transforms the diffusion distance between nodes of the graph into the Euclidean distance on the space $\mathbb{R}^\beta$. The elements $\{d_{ij}\}$ of the matrix $\mathbf{D}$ represent the distance between each pair of points x_i and x_j in the lower dimensional space $\mathbf{X} \equiv \mathbb{R}^\beta$.

Given a transition matrix $\mathbf{P}$ associated to an unbiased random walk on the graph G, we generate B bootstrap transition matrices $\{\widetilde{\mathbf{P}}_\mathbf{1}, \widetilde{\mathbf{P}}_\mathbf{2}, \ldots, \widetilde{\mathbf{P}}_\mathbf{B}\}$. We project each matrix $\widetilde{\mathbf{P}}_\mathbf{b}$ into $\mathbb{R}^\beta$ using Eq. (2), and we estimate the corresponding bootstrap distance matrix $\widetilde{\mathbf{D}}_\mathbf{b}$, whose entry d_{ij}^b is the Euclidean distance between x_i and x_j in $\mathbb{R}^\beta$. Then, we compute the average distance matrix $\widetilde{\mathbf{D}}^* = \frac{1}{B}\sum_b \widetilde{\mathbf{D}}_\mathbf{b}$. The matrix $\widetilde{\mathbf{D}}^* = \{\widetilde{d}_{i,j}^*\}$ effectively quantifies the dissimilarity between any pair of vertices of G (the smaller the distance $\widetilde{d}_{i,j}^*$ the more similar are i and j), in terms of the average distance between their projections in $\mathbb{R}^\beta$ across several replicas of $\mathbf{P}$.

Since $\widetilde{\mathbf{D}}^*$ is a dissimilarity matrix, the optimal partition of G in non-overlapping clusters can be obtained by using any method based on the iterative aggregation of clusters having minimal dissimilarity. In particular, we adopted a hierarchical clustering method which starts by considering each node as a separate cluster, and successively merges the two clusters having minimal distance. The distance of two clusters C_p and C_q is computed as the average distance between any pair of nodes (i, j) such that $i \in C_p$ and $j \in C_q$ (also known as *average linkage*,

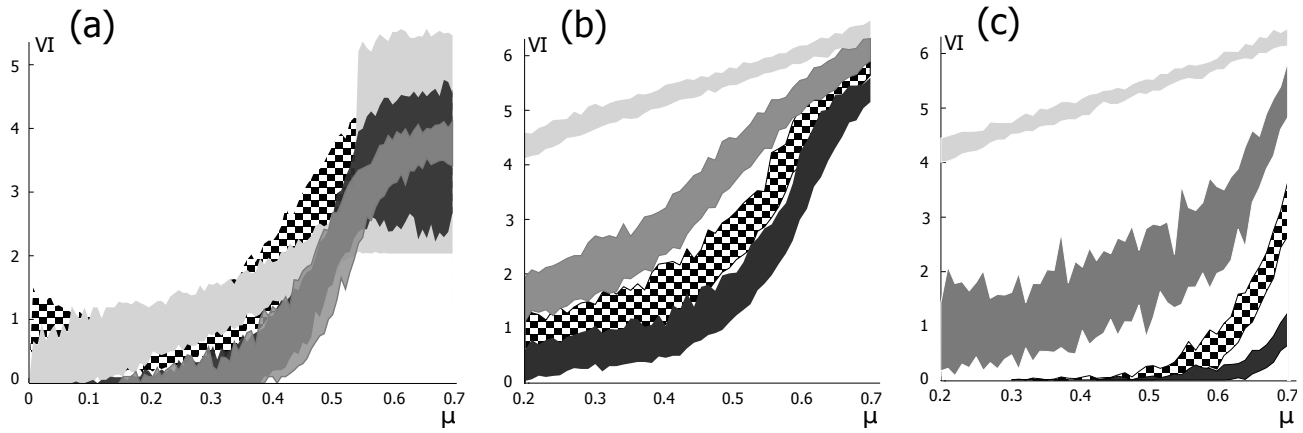

FIG. 1. **Benchmark networks.** The variation of information VI as a function of the proportion of inter-modules links μ in GN graphs (a) and as a function of the mixing parameter μ in LFR500 (b) and LFR2000 (c) graphs. The region inside each curve includes the 5^{th} and the 95^{th} percentiles of VI across R different runs. The four curves in each panel correspond to the optimal partitions obtained using the distance matrix $\mathbf{D}$ corresponding to $\mathbf{P}$ for $\beta = 1$ (light gray) and $\beta = 10$ (checked pattern), the distance matrix $\mathbf{D}^*$ corresponding to the bootstrapped matrix $\mathbf{P}^*$ for $\beta = 1$ (black), and the modularity optimization on the adjacency matrix A, as described in Refs. [7, 33] (dark-gray). The network order N, the number of runs R, and the number B of bootstrap realizations for each run are (a) $N = 128$, $R = 100$, and $B = 100$; (b) $N = 500$, $R = 100$, and $B = 100$; and (c) $N = 2000$, $R = 50$, $B = 50$.

see Sect. 4.2 in Ref.[10] for details). The algorithm stops when all the nodes have been grouped in a single cluster. The hierarchical clustering algorithm produces a dendrogram H, i.e. a tree where each of the $N - 1$ internal nodes represents the fusion of two clusters. A horizontal cut of H corresponds to a partition of the graph into a certain number of communities. The quality of each partition $\mathcal{S}$ was estimated using the modularity function [7], which compares the abundance of edges lying inside each community with respect to a null model. In formula:

$$Q(\mathcal{S}) = \sum_{s=1}^{N_s} \left[\frac{m_s}{K} - \left(\frac{k_s}{2K} \right)^2 \right], \qquad (3)$$

where N_s is the number of clusters in the partition $\mathcal{S}$, K is the total number of edges in the network, m_s is the number of edges between vertices in cluster s, and k_s is the sum of the degrees of the nodes in cluster s. The optimal partition in communities of the graph G is the cut of the dendrogram H having maximum modularity. We notice that the choice of modularity to quantify the relevance of a partition is arbitrary, and that the method proposed here can be employed also with other partition quality functions [10].

Benchmark tests.– The performance of our approach has been tested on two classes of benchmark graphs with tunable modular structure. In the first benchmark (GN), proposed by Girvan and Newman [4], each network consists of $N = 128$ nodes divided into 4 modules of equal size. Pairs of nodes in the same module are connected with probability p_{in}, while nodes belonging to different modules are linked with a probability p_{out}. Parameters are set such that the mean degree is kept constant $\langle k \rangle = 16$. By appropriately tuning p_{in} and p_{out} one can tune the percentage μ of edges lying between communities. The second class of modular graphs (LFR), proposed by Lancichinetti, Fortunato and Radicchi [31], accounts for the heterogeneity in the distributions of node degrees and community sizes. In this case, we generated modular graphs with scale-free degree distribution $P(k) \sim k^{-\gamma}$ and community size distribution $P(s) \sim s^{-\eta}$, where $\gamma = 2$ and $\eta = 1$. An appropriate tuning of the model parameters allows to create graphs with a prescribed fraction μ of inter-community edges. We considered graphs having $N = 500$ and $\langle k \rangle = 7$ (LFR500) and graphs with $N = 2000$ nodes and $\langle k \rangle = 28$ (LFR2000). The average degree of each graph was tuned to maintain the global edge density.

To compare the optimal partition produced by our bootstrap-based algorithm with the reference partition, we make use of an information theoretic measure: the variation of information (VI) [32]. In a nutshell, this non-negative metric quantifies how much information is lost and gained in changing from a partition A to a partition B. It can be estimated by $V(A,B) = -\sum_i^{c^A} \sum_j^{c^B} (\frac{n_{ij}^{AB}}{N}) \log \frac{n_{ij}^{AB}}{N} + \frac{n_{ij}^{AB}}{N} \log \frac{AB_{ij}/N}{n_i^A n_j^B/N^2}$, where c^A (c^B) is the total number of clusters in the partition A (B), n_i^A (n_j^B) is the number of nodes in the i^{th} (j^{th}) cluster of partition A (B), and n_{ij}^{AB} is the number of nodes shared by the i^{th} cluster of partition A and the j^{th} cluster of partition B. Values of V range from 0, when A and B are identical partitions, to $\log N$ when the two partitions are randomly drawn.

Fig. 1 shows the variation of information as a function of the fraction of inter-community edges μ. The reported results suggest that even when the graphs do not have any more a strong community structure, i.e. when each

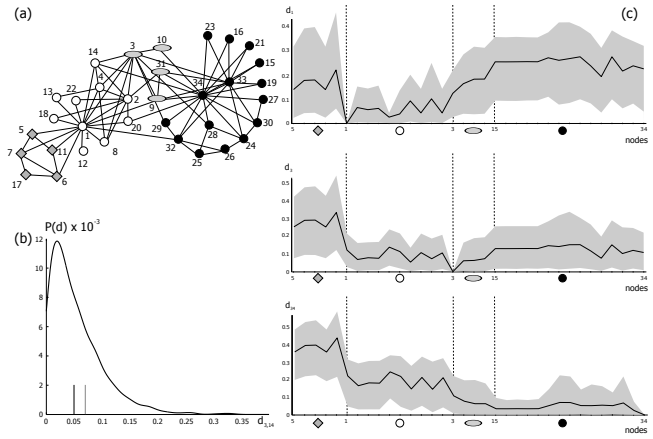

FIG. 2. (a) The best partition of the Zachary Karate club network obtained with non-parametric bootstrap for a $Q = 0.389$ (each module is represented by a different color/symbol); (b) bootstrap distribution for the distance between nodes $i = 3$ and $j = 14$: the small vertical bars indicate the value of the observed real network (gray) and the 50^{th} percentile of the distribution (black), respectively; (c) confidence intervals for the spectral distance between node $i = 1$, (top) $i = 3$ (middle), $i = 34$ (bottom) and all the other nodes rest of the nodes. Gray regions indicate the 0.05^{th}–95^{th} percentiles interval of the bootstrap distribution. Notice that node 1 is much closer to the the nodes in its community (white circles) than to the other nodes. Similarly, node 34 is very close to the other nodes in its community (black circles), but quite distant from the rest of the graph. Finally, the distance between node 3 and any node in the black community is comparable to that between node 3 and the white communities.

node has more neighbours outside its community than inside it, the accuracy of the proposed bootstrap-based method remains pretty high. For GN networks, the accuracy of the bootstrap-based method is comparable to that of a standard modularity optimization algorithm [7, 33]. For LFR500 and LFR2000, the non-parametric bootstrap method outperforms the other algorithms, even when we consider an embedding with $\beta = 1$, and exhibits a smaller value of VI up to relatively large values of μ.

Testing on real networks.– The social network of friendships between members of a karate club at a US university, studied by the anthropologist Wayne Zachary in the 1970s [34], is a paradigmatic example of network with a strong modular structure. This network ($N = 34$ nodes and $K = 78$ edges) has been considered as a benchmark reference for most of the community detection algorithms proposed in the last decade. Fig. 2(a) shows the communities detected by the bootstrap-based algorithm in the Zachary's network. It is worth noticing that the proposed method detects the two main modules (black and white circles, respectively) and an interface community between them which contains also nodes 3, 9 and 10, three nodes whose belonging to each of the two main modules has been considered unstable [15] and which have been ambiguously classified by different algorithms [35, 36].

Variability of node attributes.– Bootstrap replicates can provide useful insight about the statistical significance of any structural property of a graph with unknown and unobservable distribution. For example, in Fig. 2(b) we report the bootstrap distribution of the spectral distance between two nodes of the Karate Club network, namely, node 3 and node 14, together with the mean value (vertical gray line) and the median (vertical black line). Similarly, it is possible to compute the bootstrap confidence intervals (c.i.) of a given parameter θ as $\theta \in \left[\widehat{\theta}_b(\frac{\alpha}{2}), \widehat{\theta}_b(1-\frac{\alpha}{2})\right]$, where $\widehat{\theta}_b(\alpha)$ is the 100α percentile of the bootstrap distribution of $\widehat{\theta}$. For instance, if we consider a node i and compute the spectral distance between i and all the other nodes in the network, we can associate a confidence interval to each of these distances. In Fig. 2(c) we report the 90% confidence intervals (shaded gray) for the distance in the embedded Euclidean space $\mathbb{R}^\beta$ (here $\beta = 1$, but qualitatively similar results are obtained for $\beta = 5$ and $\beta = 10$) between node $i = 1$ (top), $i = 3$ (middle) and $i = 34$ (bottom) and the rest of the network. The solid black line represents the average over the set of replicates. Notice that node 1 and node 34 exhibit a remarkable lower distance to the other nodes of the community to which they respectively belong. Conversely, node 3 has a similar spectral distance to both the black and the white communities. This is probably the reason why this node is usually misclassified by most community detection algorithms.

Concluding remarks.– The present study combines probability theory and network science to assess the statistical significance of structural network properties by means of a non-parametric resampling scheme. We propose a method to detect communities based on the generation of bootstrap replicates of the transition matrix associated with an unbiased random walk on the graph, and we show that this method outperforms other community detection algorithms based on the spectral properties of the graph. Our results suggest that non-parametric bootstrapping of random walks can be of practical relevance for the study of the meso-scale structure of complex networks. The non-parametric bootstrapping method described here could be also valuable to assess the statistical significance of nodes attributes on the basis of other random walk parameters (e.g. hitting times), or for the statistical validation of other structural properties defined on different flavors of random walks [37, 38].

This work was supported by the EU-LASAGNE Project, Contract No.318132 (STREP). F. De Vico Fallani is financially supported by the French program "Investissements d'avenir" ANR-10-IAIHU-06.

[1] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, Phys. Rep. **424**, 175 (2006).

[2] A. Barrat, M. Barthlemy and A. Vespignani, *Dynamical processes on complex networks* Cambridge University Press, Cambridge (2008).

[3] M. Newman, *Networks: an introduction*, Oxford University Press, Oxford (2010).

[4] M. Girvan and M. E. J. Newman, *Proc. Natl. Acad. Sci. USA* **99**, 7821-7826 (2002).

[5] M. Chavez, M. Valencia, V. Navarro, V. Latora, and J. Martinerie, *Phys. Rev. Lett.* **104**, 118701 (2010).

[6] E. Bullmore and O. Sporns, *Nat. Rev. Neurosci.* **10**, 186–198 (2009).

[7] M. E. J. Newman and M. Girvan, *Phys. Rev. E* **69,** 026113 (2004).

[8] P.F. Jonsson, T. Cavanna, D. Zicha, P.A. Bates, *BMC Bioinf.* **7,** 2 (2006).

[9] R. Guimerà and L. A. N. Amaral, **Nature 433**, 895 (2005).

[10] S. Fortunato, *Phys. Rep.* **486**, 75 (2010).

[11] S. Wasserman and K. Faust, *Social Network Analysis.* Cambridge University Press, Cambridge, (1994).

[12] S. Maslov and K. Sneppen, *Science* **296**, 910 (2002).

[13] E. Ziv, R. Koytcheff, M. Middendorf and C. Wiggins, *Phys. Rev. E.* **71**, 016110 (2005)

[14] G. Bianconi, P. Pin and M. Marsili, *Proc. Natl. Acad. Sci. USA* **106**, 11433 (2009).

[15] D. Gfeller, J. C. Chappelier and P. De Los Rios, *Phys. Rev. E.* **72**, 056135 (2005)

[16] B. Karrer, E. Levina and M.E.J. Newman, *Phys. Rev. E.* **77**, 046119 (2008)

[17] M. Rosvall and C. T. Bergstrom, *PLoS ONE* **5**, e8694 (2010)

[18] A. Mirshahvalad, O. H. Beauchesne, É. Archambault and M. Rosvall, *PLoS ONE* **8**, e53943 (2013)

[19] B. Efron and R. J. Tibshirani, *An Introduction to the Bootstrap.* Chapman and Hall/CRC, New York (1993).

[20] B. Efron and R. Tibshirani, *Statist. Sci.* **1**, 54 (1986)

[21] A. Strehl, and J. Ghosh, *J. Mach. Learn. Res.* **3**, 583 (2002)

[22] Kwak, H., Choi, Y., Eom, Y.-H., Jeong, H., and Moon, S., *Proceedings of IMC '09*, 301–314 (2009).

[23] E.-Y. Kim, D.-U. Hwang, and T.-W. Ko, *Phys. Rev. E.* **85**, 026119 (2012).

[24] A. Lancichinetti and S. Fortunato *Sci. Rep.* **2**, 336 (2012);

[25] D. M. Cvetkovic, M. Doob, and H. Sachs, *Spectra of Graphs: Theory and Applications.* Johann Ambrosius Barth Verlag, Heidelberg, (1995).

[26] F. R. K. Chung, *Spectral Graph Theory.* American Mathematical Society, Providence, RI, (1997).

[27] I. V. Basawa, T. A. Green, W. P. McCormick, and R. L. Taylor, *Commun. Stat.-Theory Methods* **19**, 1493 (1990)

[28] P. P. Sendi, H. C. Bucher, B. A. Craig, D. Pfluger, and M. Battegay, *J. Acquir. Immune Defic. Syndr. Hum. Retrovirol.* **20**, 376 (1999)

[29] D. Gfeller, and P. De Los Rios, *Phys. Rev. Lett.* **99**, 038701 (2007).

[30] R. R. Coifman and S. Lafon, *Appl. Comput. Harmon. Anal.* **21** 5 (2006).

[31] A. Lancichinetti, S. Fortunato and F. Radicchi, *Phys. Rev. E.* **78**, 046110 (2008)

[32] M. Meilă, *J. Multivariate Anal.* **98**, 873 (2007).

[33] M. E. J. Newman, *Proc. Natl. Acad. Sci. USA* **103**, 8577 (2006).

[34] W. W. Zachary, *J. Anthropol. Res.* **33,** 452-473 (1977).

[35] D. Li, I. Leyva, J.A. Almendral, I. Sendiña-Nadal, J.M. Buldú, S. Havlin and S. Boccaletti, *Phys. Rev. Lett.* **101**, 168701 (2008).

[36] E. Estrada, *Chaos* **21**, 016103 (2011).

[37] C. Allefeld and S. Bialonski, *Phys. Rev. E.* **76**, 066207 (2007);

[38] R. Sinatra, D. Condorelli, and V. Latora, *Phys. Rev. Lett.* **105**, 178702 (2010).